(12) UK Patent Application (19) GB (11) 2 345 783 (13) A

(43) Date of A Publication 19.07.2000

(21) Application No 0000558.7

(22) Date of Filing 11.01.2000

(30) Priority Data
(31) 9900609 (32) 12.01.1999 (33) GB

(71) Applicant(s)
The Speech Recognition Company
(Incorporated in the United Kingdom)
16/18 Empress Road, LONDON, SW6 1TT,
United Kingdom

(72) Inventor(s)
Mike Craig
Anthony Buxton

(74) Agent and/or Address for Service
Stevens Hewlett & Perkins
Halton House, 20/23 Holborn, LONDON, EC1N 2JD,
United Kingdom

(51) INT CL[7]
G10L 15/26 , G06F 3/16 // G10L 15:14 101:023 101:027

(52) UK CL (Edition R)
G4R RHB RRL R1F
U1S S2125 S2127

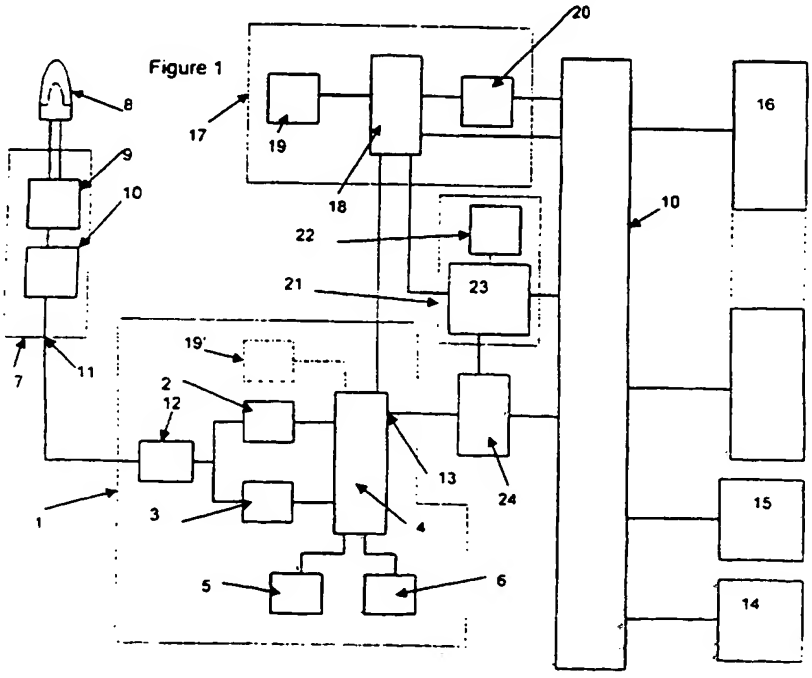
(56) Documents Cited
US 5465378 A

(58) Field of Search
UK CL (Edition R) G4R REX RHB RRL
INT CL[7] G06F 3/00 3/16 , G10L 15/00 15/06 15/08
15/10 15/14 15/22 15/26
Online:WPI, EPODOC, JAPIO

(54) Abstract Title
Speech recognition system

(57) A speech recognition system consists of a conventional speech recognition device (1) for receiving speech data via an interface (7) with a microphone (8). The speech recognition device (1) includes a dictation grammar (5) and a likelihood analyser (4) for identifying the most likely words uttered by a user into the microphone. In addition to the conventional speech recognition device (1) a speech recognition adapter (17) is provided that has an adaptive memory (19) in which is stored and developed new dictation grammars and a data builder (18). The speech recognition adapter (17) enables the speech recognition system to be used in circumstances where there is little, if any, textual context to the dictation - for example where data is being dictated into data records or forms.

Figure 1



GB 2 345 783 A

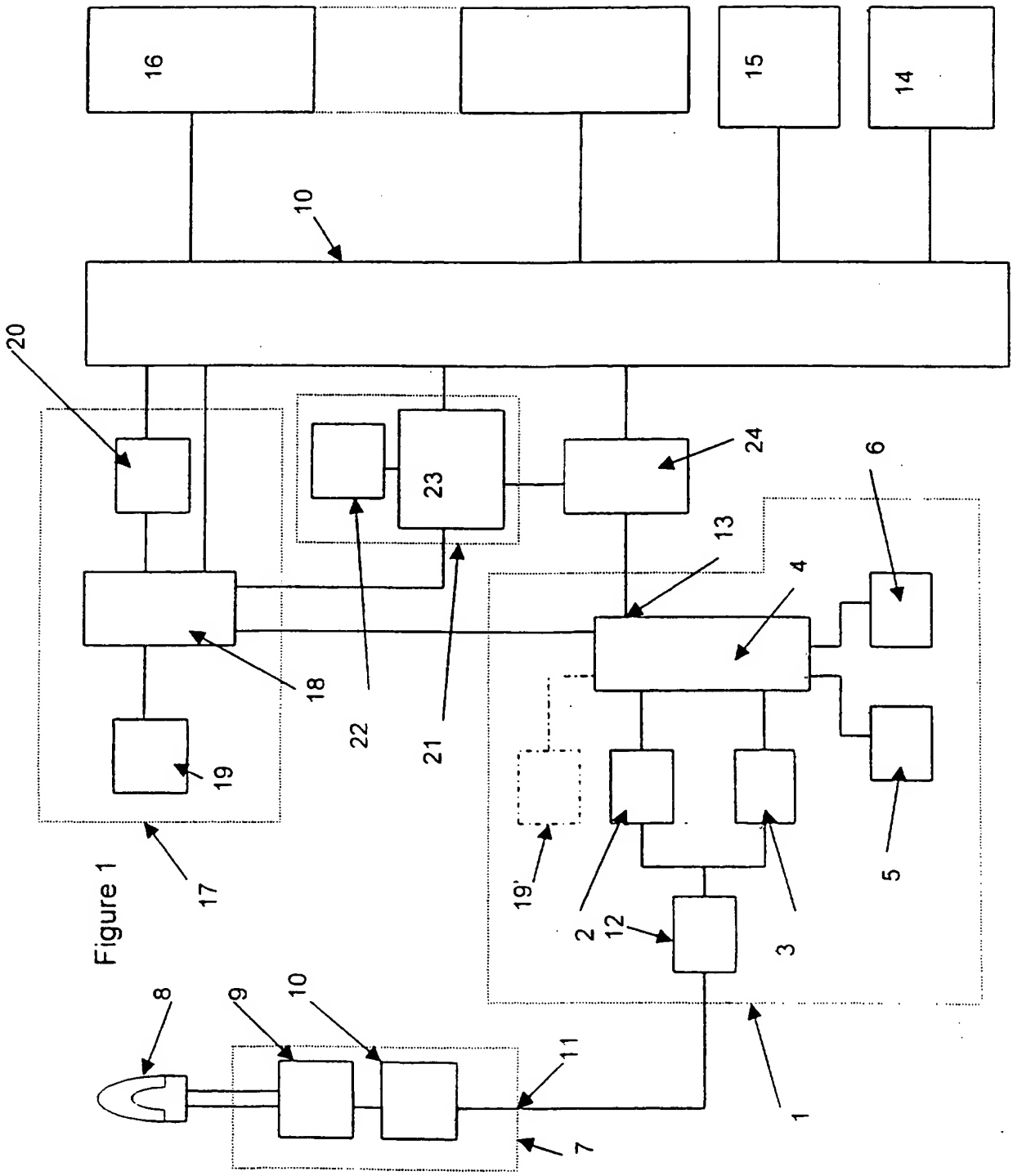


Figure 1

SPEECH RECOGNITION SYSTEM

The present invention relates to a speech recognition system. More particularly the present invention relates to a speech recognition dictation system suited for use with databases and non-word processing software packages.

Existing speech recognition systems that use dictation grammars employ Markov models to recognise dictated speech from the viewpoint of probability. An example of such a system may be found in EP-A-033067 which describes in particular means whereby the vector quantisation code book inherent in the Markov model may be adapted to accommodate different speakers and/or different environments. In order to recognise words uttered by a speaker, the speech recognition system considers the context of the adjacent words and predicts, using probabilistic analysis, the most likely combination of words matching the recorded speech. Thus, the accuracy of such speech recognition systems is partially dependent on the textual context of the utterances of the speaker. This in turn means that the performance of such systems is most reliable for applications in which text documents are created containing context sensitive streams of information. In the absence of context sensitive streams of information the performance of speech recognition systems can be poor. For example where information is to be dictated into data records or into input forms.

Although some existing speech recognition systems permit their use with form applications, in general each field of the form must be separately identified in turn with the data to be entered either by means of mouse clicks, key-strokes or verbal commands. In any event the performance of speech recognition systems in recognising the dictated field entries remains poor as the conventional probabilistic analysis cannot be employed in the absence of any adjacent words to provide context.

The basic dictation grammars of most current speech recognition dictation systems are based on the contents of one or more selected newspapers. The probability analysis performed during recognition of the

dictation is thus partially based on how often individual words appeared in the newspapers. For more specialist needs add-on dictation grammars or vocabularies have been developed, for example for legal or medical use. However, even with these add-on vocabularies the recognition of individual dictated words employs the same probabilistic analysis based on the context of adjacent words. Hence, even with such add-on vocabularies, current speech recognition systems are not adapted to perform well with databases or other applications that do not involve the dictation of a stream of words having a common textual context.

The present invention seeks to provide an improved speech recognition system suitable for dictation into databases and other software applications where substantially no textual context is available for probabilistic analysis. The present invention also seeks to provide separately a speech recognition adapter that may be used to adapt a conventional speech recognition system to render the system suitable for use with databases and other software applications where substantially no textual context is available.

The present invention provides a speech recognition system for use with a software application having a plurality of individual data entry domains, the speech recognition system including:

an input for receiving an input signal representing dictated text intended for completion of at least one data entry domain of a software application;

a spectral analyser for analysing the input signal;

a likelihood analyser in communication with the spectral analyser for matching the input signal with one or more stored words; and

an output for supplying the most likely word or words to have been dictated corresponding to the one or more stored words matched by the likelihood analyser,

wherein the speech recognition system further includes:

an interrogator for addressing and analysing the software application to extract information at least on relationships between the individual data

entry domains of the software application and the data contained therein;

an adaptive memory in which is stored a lexicon containing words suitable for entry into the data entry domains, each of the words having assigned weighing values; and

a data builder in communication with the interrogator and the adaptive memory for determining in dependence on the output of the interrogator suitable words and their weighting values for the lexicon and wherein the likelihood analyser is in communication with the adaptive memory whereby the likelihood analyser is able to match the input signal with stored words in the lexicon in dependence on the weighting values.

In an alternative aspect the present invention provides a speech recognition adapter comprising:

an input for communication with a software application having a plurality of individual data entry domains;

an interrogator for addressing and analysing the software application to extract information on relationships between the individual data entry domains of the software application and the data contained therein;

an adaptive memory in which is stored a lexicon containing words suitable for entry into the data entry domains, each of the words having assigned weighing values;

a data builder in communication with the interrogator and the memory for determining from the output of the interrogator suitable words and their weightings for the lexicon; and

an output for communication with a speech recognition device whereby the speech recognition device is able to access the adaptive memory and match input dictated text with one or more stored words in the lexicon in dependence on the weighting values.

In a still further aspect the present invention provides a speech recognition method for identifying dictated text intended for insertion into one or more data entry domains of a software application, the method including:

interrogating the software application to extract information at least

on relationships between the individual data entry domains of the software application and the data contained therein;

generating a lexicon containing stored words suitable for entry into the data entry domains;

assigning, in dependence on the results of the interrogation, suitable weighting values to each of the stored words of the lexicon;

spectrally analysing an input signal representing the dictated text;

matching the input signal with one or more words stored in the lexicon in dependence on the weighting values; and

outputting the most likely word or words to have been dictated corresponding to the one or more stored words with which the input signal was matched.

It will of course be understood that in the context of this document reference to words is intended to encompass all intelligible utterances and in particular numerals and letters.

An embodiment of the present invention will now be described by way of example, with reference to Figure 1 which is a schematic diagram showing a speech recognition system in accordance with the present invention.

The speech recognition system shown in Figure 1 includes a conventional speech recognition device 1 that includes an utterance detector 2 for identifying when an input represents dictation rather than noise; a frame buffer 3 for storing a series of short time segments of an input signal; a likelihood analyser 4 for calculating likelihood scores representing the probabilities of a given segment matching part of one or other words; a dictation grammar or vocabulary memory 5; and a potential match memory 6. The vocabulary memory 5 contains data on the vocal characteristics of one or more separately identified users along with a large vocabulary of words, for example 30,000 words sometimes as many as 60,000 words, each of which is weighted with respect to its likelihood of occurrence whereas the potential match memory 6 is a random access memory in which is temporarily stored a selection of words from the

vocabulary memory 5 that represent potential matches for one or more sequential segments of the input signal. Such a conventional speech recognition apparatus is described in US 4783803 the contents of which is herein incorporated by reference.

The speech recognition device 1 receives speech data, via an interface 7, from a microphone 8 into which input analog speech data is dictated. The interface 7 includes an amplifier 9 and an analog-to-digital converter 10 where the input analog speech data is converted to digital data. The output 11 of the interface 7 is connected to the speech recognition device 1 and in particular to the input of a spectral analyser in the form of a fast Fourier Transform (FFT) device 12 that converts the input digital data from the time domain to the frequency domain. That is the input digital data is divided into short time segments, for example each segment may relate to a 0.02 s section of the digital data and each segment is analysed to determine the energy amplitude of the recorded dictation at a plurality of different discrete frequencies. In an idealised speech recognition system the spectral analysis of the individual time segments would enable identification of the individual phonemes of the dictated words. Such a procedure though is memory hungry and so in practice pattern matching of the Fourier Transforms is performed by the likelihood analyser 4 to identify the most likely words to have been dictated.

The output of the Fourier Transform circuit 12 is connected to the utterance detector 2 and the frame buffer 3 which ideally is large enough to hold enough segments that would normally be expected to be contained in a word. If the input signal is deemed to be dictation rather than noise by the utterance detector 2 the likelihood analyser 4 is enabled and the contents of the frame buffer 3 is compared with the contents of the vocabulary memory 5 in a rapid match computation. This comparison enables the likelihood analyser 4 to identify a smaller group of words in the vocabulary that are potential matches. This smaller group is temporarily stored in the potential match memory 6 whilst the likelihood analyser 4 performs a more detailed comparison to identify the word having the best

match.

The output 13 of the speech recognition device 1 is in communication with the systems platform 10, which in many cases is likely to be a Windows™ environment. In this way text generated by the speech recognition device 1 is input into one or other applications or databases 16 via the systems platform 10. An output device 14 such as a conventional monitor and/or printer port is also provided along with a user input terminal 15 such as a conventional keyboard and/or mouse.

For conventional speech recognition, the likelihood analyser 4 of the speech recognition device 1 considers the likelihood or probability of each potential word preceding or following other potential words using HMM (hidden Markov modelling) and the weightings of the individual words of the dictation grammar reflect these probabilities. However, this analysis performs poorly when there is little or substantially no textual context to the dictated words. To improve the performance of the speech recognition device 1 when a user is dictating data having little or no common textual context into a database or other software application, a speech recognition adapter 17 is provided. The speech recognition adapter 17 has a data builder 18 that includes an adaptive memory 19 in which one or more new dictation grammars/vocabularies can be stored and developed. The adaptive memory 19 will usually contain words already stored in the vocabulary memory 5 of the speech recognition device, in many cases with different weightings, along with new words with their assigned weightings. The data builder 18 is in communication with the systems platform 10 so that a user may provide additional information to the data builder directly using the input terminal 15.

As shown in Figure 1 the data builder 18 communicates with the likelihood analyser 4 of the speech recognition device 1 enabling access to the adaptive memory 19 of the adapter 17 to the likelihood analyser 4. Alternatively the contents of the adaptive memory 19 may be transferred from the adapter 17 to the speech recognition device 1 and stored in a separate memory 19' (dotted lines in Figure 1) or may replace the

vocabulary memory 5. It will be clear that in such circumstances the likelihood analyser 4 addresses the adaptive memory directly.

The speech recognition adapter 17 also includes an interrogator 20 that is connected to the systems platform 10 and to the data builder 18. The interrogator 20 is in the form of a data mining or report writing program and is generally conventional in nature. It is adapted to address databases and other software applications via the systems platform 10 to extract information and data relationships contained within the applications. Data mining is a conventional software tool that has been developed in recent years, primarily for use in marketing analysis such as consumer surveys and also on the vast quantities of data now available on the internet, to enable very large amounts of data to be analysed extremely efficiently to extract useful information. The interrogator 20 is used to identify the frequency of individual words specific to an existing database or other software application, to identify relationships between individual fields in a database and the entries made in that field and to identify relationships between entries in different fields. For example the interrogator 20 may identify a link such that of City=London and Postcode starts SW10, then Telephone Number will start with 0171 352. The information extracted by the interrogator 20 is used by the data builder 18 to adjust the weightings of words stored in the adaptive memory 19 so that the weightings reflect the specific features of the database or application. In other words, although the contents of a database may contain little or no textual context sensitivity, field context sensitivity is identified by the interrogator and is reflected in the weightings of the contents of the adaptive memory 19 which is in the form of a database grammar.

The speech recognition adapter 17 enables improved word recognition by the speech recognition device 1 through tailored weightings of individual words to reflect the field context and field data sensitivity of a particular existing database or application. However, the interrogator 20 is passive in that it analyses the contents of a database or other application, the interrogator 20 generally does not provide information on how the

database is habitually updated, for example the order in which the fields or domains of the database are populated. The speech recognition system therefore additionally includes a domain analyser 21 that monitors a database or other application 16 via the systems platform 10 when the database is being updated. The domain analyser 21 includes a domain route memory 22 and a domain route predictor 23. The domain route predictor 23 monitors the order in which each domain of a particular database, for example, is in turn updated by a user by means of the input terminal 15, a mouse or through dictated entries. Where patterns of behaviour or habitual routes are identified these are stored in the domain route memory 22. The patterns may additionally be communicated to the data builder 18 so that the weightings of words in the adaptive memory may be altered to reflect the likelihood of particular words, or keystrokes, following one another. Hence, movements through a database can become implicit, rather than explicit. In this way the data builder 18 can provide weightings for the most likely next domain once data has been dictated in an earlier domain and also the most likely next data entry to be made. The patterns stored in the domain route memory 22 may be specific to an identified user which in turn can result in the weightings of the words in the adaptive memory also being specific to a particular user.

The domain analyser 21 is also in communication with a data entry monitor 24 that also receives the output from the speech recognition device 1. In most cases the output of the speech recognition device 1 is fed without interruption by the data entry monitor 24 to the systems platform 10 and thence to the particular database 16 currently in use. However, where habitual routes through the database have been stored in the domain route memory 22, the data entry monitor 24 may interrupt the output of the speech recognition device to direct the output of the speech recognition device to the most likely next domain without the need for the user to dictate such a movement. Where the user dictates data for more than one domain in a single utterance, the data entry monitor may interrupt the data to direct different portions of the output of the speech recognition device to

one or more separate domains. Furthermore, where the output of the speech recognition device 1 is deemed inappropriate for the most likely next domain, the data entry monitor 24 may issue a query to the user either via the monitor 14 or an audible message or beep questioning the accuracy of the output of the speech recognition device or the domain for which the dictated data is intended.

All devices shown in Figure 1 with the exception of the microphone 8, the amplifier 9, the display device 14 and the input terminal 15 are provided in a workstation in the form of software. The speech recognition system may be implemented on a conventional PC, having for example a 133 MHz Pentium ™ processor, an industry standard 16-bit sound card such as Creative Labs ™ Sound Blaster ™ 16 sound card and upwards of 64 MB. Preferably, the PC is operated through a Windows 95 ™ or Windows NT ™ environment. Alternatively, the speech recognition system may be implemented in dedicated hardware.

Ideally, when in use the domain analyser 21 is transparent to the user. The database is displayed on the monitor 14 and the domain adapter 21 works in the background to move the cursor automatically about the various domains or fields in the database. The order in which the domains are selected in turn may be fixed and pre-programmed by the user. Alternatively, through monitoring the habitual order in which a user completes each of the fields, the domain analyser 21 may predict the most likely next field. Thus, when the domain analyser 21 is enabled, preferably its existence is only signified by the presence of a minimised icon on the output monitor 14. This may also be true for the speech recognition adapter 17. When the adapter 17 is implemented as part of a speech recognition system, the adapter 17 provides a lexicon for use by the speech recognition device 1 specific to the particular database or software application that is being run on the system. Hence, here too the adapter 17 may be transparent to the user and may be represented simply by a minimised icon. Naturally, the speech recognition device 1 can be used in its conventional mode for word processing. Although the speech

recognition adapter 17 and the domain analyser 21 are shown in Figure 1 separately it will be understood that they may be combined as part of a single utility.

When the speech recognition adapter 17 is enabled and a database 16 or other software application is accessed by a user, the interrogator 20 addresses the database to determine the different fields or domains to which data may be added, the size and any specific characteristics of the domains, for example one or more of the domains may only accept numerals. Some of the domains may have indexes associated with them and these indexes are also transferred to the data builder 18 and to the adaptive memory 19. The data builder 18 assigns a plurality of weightings to each data entry in the adaptive memory 19 with respect to each of the individual domains, to the domain names themselves and to data entries in other domains. The initial weightings of the individual domain names may be determined by the order in which they appear in the application. Alternatively, default weightings may be employed or as mentioned above the user may pre-program the desired order in which the domains are to be addressed.

Where the speech recognition analyser 17 is used to produce a database grammar or vocabulary for an existing database the analysis of the database by the interrogator 20 may be performed only once with the vocabulary produced by the data builder subsequently being input to the speech recognition device either in a separate memory 19' or as a replacement for the contents of the memory 5. Even when the speech recognition analyser is implemented as part of a speech recognition system, preferably the interrogator 20 would not be enabled every time the database is accessed by the user. Instead, the interrogator 20 ideally would be enabled at predetermined periods, for example once a month, to refresh the existing adaptive memory 19 for that particular database.

When the domain analyser 21 is enabled and a new database is accessed by the user, the domain analyser 21 monitors the order in which the individual domains of the database are completed by the user and

identifies the habitual order adopted by the user for completing the database entries and stores this in the domain route memory 22. When an existing database is accessed by a user, the domain analyser 21 determines from the domain route memory 22 the most likely first domain to be completed by the user. Once the most likely domain has been identified, the domain analyser 21 communicates to the data entry monitor 24 the most likely domain which then causes the most likely domain to be automatically selected. In this way the user may dictate the data to be entered into that domain without first separately identifying the domain. The data dictated by the user is then analysed by the speech recognition device 1 using the specific vocabulary generated by the speech recognition adapter 17. The speech recognition device 1 preferably employs hidden Markov probabilistic analysis to identify the most likely word(s) to have been dictated by the user and outputs the word(s) via the data entry monitor to the systems platform 10. Once the data has been entered into the domain, the domain analyser 21 determines the most likely next domain to be completed and accordingly communicates this to the data entry monitor 24. The process is then continued as the database entry in completed. If at any time the user selects a domain different to the predicted most likely next domain the domain analyser 21 can adjust the weightings of the remaining domains accordingly. The same is true of the weightings of the words in the adaptive memory 19 which are adjusted by the speech recognition device 1 to reflect the changed order in which the data is being entered by the user. Also, if the user dictates data that is not permissible for the next most likely domain, for example the domain is limited to numerical entries and the user has dictated a name, the data entry monitor 24 issues a query to be displayed in the monitor 14 asking the user to specify the correct domain.

The same process is cyclically repeated for each domain of the database in turn as the user dictates new data and for each new entry. In each case the weightings of the domain names and data entries stored in the adaptive memory 19 are altered to reflect the domains and data already

recognised by the system. This results in the accuracy and speed of the predictions increasing as a user adds more data.

To assist in an understanding the following is an example of how the speech recognition system may interact with an address database. The address database may have ten separate domains as follows: Title, First Name, Second Name, Surname, House Number, Street Name, Town, County, Postcode and Telephone Number. The Title domain may have an index associated with it listing "Mr, Ms, Mrs, Miss, Dr". The Postcode domain may have an associated rule that the data must form an alphanumeric string and the Telephone Number domain may have the rule that relevant data entries must be purely numerical.

In addition to the data entry restrictions identified above, the interrogator 20 may identify additional rules concerning relationships between data in separate domains of the database. For example, the weightings for certain numerical strings, with respect to the Telephone Number domain, may be adjusted in dependence on the particular data entry in the Town domain. Hence, if the data entry for the Town domain is London then the weightings of the data entries 0171 or 0181 are greatly increased.

When the address database is first opened the most likely first domain i.e. the domain having the greatest weighting, is the Title domain. However, the user dictates "Mr John F Brown". Even though the dictation is made as a single string, because the string is too long for only the Title domain and the weightings of the various name domains are also high, the data entry monitor 24 determines that the output of the speech recognition device 1 must cover data entries for more than one domain. The data entry monitor 24 therefore determines from the domain analyser 21 that the next most likely domains are the Name domains. The data entry monitor 24 therefore interrupts the output of the speech recognition device 1 to direct different portions of the output to the Title domain and then to the different Name domains in turn.

Once the Name domains have been completed, the House Number

domain is identified by the domain analyser 21 as the next most likely domain and in the adaptive memory 19 the weightings of individual numerals is high with respect to having been preceded by data for the Name domain. Following dictation and completion of the House Number domain, the Street Name and Town domains etc are completed in due course in a similar manner.

Finally, only one domain remains, the Telephone Number domain. However, the user is interrupted and decides to save the data already dictated. The user therefore dictates a global command rather than the predicted numerical string for the Telephone Number domain. The weightings of the global commands stored in the vocabulary memory 5 are high and so, for each new segment of dictation, the probability of the dictation containing one or more global commands is assessed by the likelihood analyser 4. In this example, the global command made by the user to save the data already dictated would be recognised as the global command and appropriate action instructed through the systems platform 10.

Where a selected domain has rules restricting the type of data that can be entered into the domain for example numerical data for the Telephone Number domain, dictation that does not accord to the rules for that domain may be rejected by the database and a query generated on the monitor 14. Thus, in the case of the Telephone Number domain, if the user has dictated non-numerical data, the database can prompt the user for dictation specific to the Telephone Number domain. Also if there is a Fax Number domain as well as a Telephone Number domain, the speech recognition device 1 may include macros for natural languages commands such as "repeat" or "fax number equals telephone number" to avoid the need for the user to repeat the data a second time.

The speech recognition device 1 may also include macros that automatically link to other software packages in response to global commands. For example, where a user dictates "3600 divide by 12" the speech recognition device identifies the dictation as being reference to a

mathematical procedure and automatically opens an maths software package to run the calculation and insert into the selected domain of the database, not the dictated calculation, but the total calculated by the maths software program.

Although reference has been made herein to Markov probabilistic techniques, it will be appreciated that alternative probabilistic analyses may be adopted, where appropriate. It should be understood that the speech recognition system is intended for use with all main database types including Dbase™ and Access™ etc. Naturally, the speech recognition system is intended for multiple language support including UK English, US English or French etc.

The speech recognition system described above is particularly suited for use with databases and other non-word processing applications where there is limited, if any, textual context to aid in providing an accurate recognition of the words dictated by the user. Contrary to conventional speech recognition, the speech recognition system employs domain and data entry context to provide an accurate recognition of the dictated words instead of textual context. Hence, the weightings of the data entries and the individual words in the data entries for any particular domain are determined in dependence on that domain and the data entries in related domains. Thus, the adaptive memory 19 may be described as containing a database grammar rather than the conventional dictation grammars used in existing dictation systems.

Textual context may still be employed by the speech recognition device 1 as appropriate. For example, when the database or application includes a free text field where context sensitive text is expected to be dictated by the user. In the free text field the speech recognition device's own vocabulary memory 5, rather than the adaptive memory 19, can be employed as the resource for recognition purposes.

The speech recognition system provides the advantage of being able to anticipate the next most likely utterances of the user and thereby improve the speed and accuracy of recognition. Also, the speech

recognition system is capable of self-learning. It interrogates the database to identify the domains for which data entries are required, the system also adaptively weights individual data entries in dependence on the identified domains and on the habits of the user.

The speech recognition system also incorporates the same user friendly error correction facilities as are available with existing speech recognition systems and enables the weightings to be adjusted accordingly. Thus, a user can both play back his own dictation or have the speech recognition system say back what it recognised from the dictation. For example, if the user were to dictate "London" for a city domain but speech recognition system recognised "Luton", the system could be set up by the user to say back each entry once recognised so that errors in recognition can be quickly identified. The speech recognition system may additionally provide prompts to the user with information on the appropriate commands for the user to give at that stage in the running of the application. This can be conceived as "you can say" prompts and can also be employed where the speech recognition determines that an ambiguous command has been dictated by the user.

Thus, it may be seen that the speech recognition system described above has greater flexibility and is more powerful than conventional speech recognition systems as it incorporates all of the features of existing systems and adds to that the ability to recognise with higher accuracy dictation where little or no textual context is available by analysing the domain context of the dictated data. With the speech recognition system of the present invention higher accuracy dictation into database applications and other non-word processing applications is now possible.

As shown in Figure 1 the speech recognition system includes the speech recognition device, the speech recognition adapter and the domain analyser. However, existing conventional speech recognition systems may be updated to enable such conventional systems to perform reliably with non-text sensitive dictation. The speech recognition analyser and the domain analyser may be separate and can be used to generate a weighted

vocabulary specific to a particular existing database in which the weightings of the words in the vocabulary are dictated both by how often they appear in the database but also the order in which the domains are likely to be completed. The weighted vocabulary may then be loaded as a separate adaptive memory to the conventional speech recognition system.

CLAIMS

1. A speech recognition system for use with a software application having a plurality of individual data entry domains, the speech recognition system including:

an input for receiving an input signal representing dictated text intended for completion of at least one data entry domain of a software application;

a spectral analyser for analysing the input signal;

a likelihood analyser in communication with the spectral analyser for matching the input signal with one or more stored words; and

an output for supplying the most likely word or words to have been dictated corresponding to the one or more stored words matched by the likelihood analyser,

wherein the speech recognition system further includes:

an interrogator for addressing and analysing the software application to extract information at least on relationships between the individual data entry domains of the software application and the data contained therein;

an adaptive memory in which is stored a lexicon containing words suitable for entry into the data entry domains, each of the words having assigned weighing values; and

a data builder in communication with the interrogator and the adaptive memory for determining in dependence on the output of the interrogator suitable words and their weighting values for the lexicon and wherein the likelihood analyser is in communication with the adaptive memory whereby the likelihood analyser is able to match the input signal with stored words in the lexicon in dependence on the weighting values.

2. A speech recognition system as claimed in claim 1, wherein the interrogator additionally extracts information on at least one of data entry domain names, relationships between data entry domains, and the occurrence of data in the domains.

3. A speech recognition system as claimed in either of claims 1 or 2, wherein the interrogator is a data mining device or report writing device.

4. A speech recognition system as claimed in any one of the preceding claims, further including a domain analyser adapted to monitor the order in which individual data entry domains of the software application are populated.

5. A speech recognition system as claimed in claim 4, wherein the domain analyser includes a domain route memory for storing the frequency with which individual data entry domains are populated after other domains.

6. A speech recognition system as claimed in claim 5, wherein the domain analyser further includes a domain route predictor for identifying the most likely next domain to be populated and a domain route monitor in communication with the output of the speech recognition system for supplying instructions identifying the next domain for population.

7. A speech recognition system as claimed in any one of the preceding claims, wherein the spectral analyser is in the form of a Fast Fourier Transform device.

8. A speech recognition system as claimed in any one of the preceding claims, wherein there is further provided a vocabulary memory containing a plurality of stored words each with a respective acoustic model and associated weighting values.

9. A speech recognition system as claimed in any one of the preceding claims, wherein the stored words contained in the adaptive memory include numerals and letters.

10. A speech recognition adapter comprising:

an input for communication with a software application having a plurality of individual data entry domains;

an interrogator for addressing and analysing the software application to extract information on relationships between the individual data entry domains of the software application and the data contained therein;

an adaptive memory in which is stored a lexicon containing words suitable for entry into the data entry domains, each of the words having assigned weighing values;

a data builder in communication with the interrogator and the memory for determining from the output of the interrogator suitable words and their weightings for the lexicon; and

an output for communication with a speech recognition device whereby the speech recognition device is able to access the adaptive memory and match input dictated text with one or more stored words in the lexicon in dependence on the weighting values.

11. A speech recognition adapter as claimed in claim 10, wherein the interrogator additionally extracts information on at least one of data entry domain names, relationships between data entry domains, and the occurrence of data in the domains.

12. A speech recognition adapter as claimed in either of claims 10 or 11, wherein the interrogator is a data mining device or report writing device.

13. A speech recognition adapter as claimed in any one of claims 10 to 12, further including a domain analyser adapted to monitor the order in which individual data entry domains of the software application are populated.

14. A speech recognition adapter as claimed in claim 13, wherein the domain analyser includes a domain route memory for storing the frequency

with which individual data entry domains are populated after other domains.

15. A speech recognition adapter as claimed in claim 14, wherein the domain analyser further includes a domain route predictor for identifying the most likely next domain to be populated and a domain route monitor in communication with the software application for supplying instructions identifying the next domain for population.

16. A speech recognition adapter as claimed in any one of claims 10 to 15, wherein the adaptive memory contains acoustic models for at least some of the stored words in the lexicon.

17. A speech recognition adapter as claimed in any one claims 10 to 16, wherein the stored words contained in the adaptive memory include numerals and letters.

18. A speech recognition method for identifying dictated text intended for insertion into one or more data entry domains of a software application, the method including:

interrogating the software application to extract information at least on relationships between the individual data entry domains of the software application and the data contained therein;

generating a lexicon containing stored words suitable for entry into the data entry domains;

assigning, in dependence on the results of the interrogation, suitable weighting values to each of the stored words of the lexicon;

spectrally analysing an input signal representing the dictated text;

matching the input signal with one or more words stored in the lexicon in dependence on the weighting values; and

outputting the most likely word or words to have been dictated corresponding to the one or more stored words with which the input signal was matched.

19. A speech recognition method as claimed in claim 18, further including monitoring the order in which the data entry domains are populated.

20. A speech recognition method as claimed in claim 19, wherein the frequencies with which individual data entry domains of the software application are populated after other domains are stored.

21. A speech recognition method as claimed in claim 20, further including the step of outputting instructions for selection of the most likely next data entry domain to be populated.

The Patent Office

INVESTOR IN PEOPLE

22

Application No: GB 0000558.7
Claims searched: 1 to 21

Examiner: John Donaldson
Date of search: 22 February 2000

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.R): G4R(REX, RHB, RRL)

Int Cl (Ed.7): G06F 3/00, 3/16; G10L 15/00, 15/06, 15/08, 15/10, 15/14, 15/22, 15/26

Other: Online:WPI, EPODOC, JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	US 5465378 (DUENSING), see abstract	-

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.